Review. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are orthogonal, the orthogonal projection of $\boldsymbol{w}$ onto $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is

$$\hat{\boldsymbol{w}} = \frac{\boldsymbol{w} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1} \boldsymbol{v}_1 + \ldots + \frac{\boldsymbol{w} \cdot \boldsymbol{v}_n}{\boldsymbol{v}_n \cdot \boldsymbol{v}_n} \boldsymbol{v}_n.$$

Example 69.

(a) Project $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ onto $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \right\}$.

(b) Express $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ in terms of the basis $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$.

Solution.

(a) We note that the vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ are orthogonal to each other.

Therefore, the projection can be computed as $\frac{\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \frac{\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \frac{8}{6} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \frac{4}{5} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$.

Comment. If we didn't have an orthogonal basis for $W = \operatorname{col}\left(\begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 0 \end{bmatrix} \right)$, then we would have to solve the least squares problem $\begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ instead to get the same final result (with more work).

(b) Note that this basis is orthogonal! Therefore, we can compute $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \frac{8}{6} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \frac{4}{5} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + \frac{5}{30} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$.

(We proceed exactly as in the previous part to compute each coefficient as a quotient of dot products.)

Gram–Schmidt

(Gram–Schmidt orthogonalization)

Given a basis $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots$ for W, we produce an orthogonal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ for W as follows:

- $\boldsymbol{q}_1 = \boldsymbol{w}_1$
- $\boldsymbol{q}_2 = \boldsymbol{w}_2 - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_2 \text{ onto } \boldsymbol{q}_1 \end{pmatrix}$
- $\boldsymbol{q}_3 = \boldsymbol{w}_3 - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_1 \end{pmatrix} - \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_2 \end{pmatrix}$
- $\boldsymbol{q}_4 = \ldots$

Note. Since $\boldsymbol{q}_1, \boldsymbol{q}_2$ are orthogonal, $\begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \operatorname{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2\} \end{pmatrix} = \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_1 \end{pmatrix} + \begin{pmatrix} \text{projection of} \\ \boldsymbol{w}_3 \text{ onto } \boldsymbol{q}_2 \end{pmatrix}$.

Important comment. When working numerically on a computer it actually saves time to compute an orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ by the same approach but always normalizing each $\boldsymbol{q}_i$ along the way. The reason this saves time is that now the projections onto $\boldsymbol{q}_i$ only require a single dot product (instead of two). This is called **Gram–Schmidt orthonormalization**. When working by hand, it is usually simpler to wait until the end to normalize (so as to avoid working with square roots).

Note. When normalizing, the orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ is the unique one (up to $\pm$ signs) with the property that $\operatorname{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k\} = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ for all $k = 1, 2, \ldots$.

Armin Straub
straub@southalabama.edu

Example 70. Using Gram–Schmidt, find an orthogonal basis for $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Solution. We already have the basis $\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\boldsymbol{w}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ for W. However, that basis is not orthogonal.

We can construct an orthogonal basis $\boldsymbol{q}_1, \boldsymbol{q}_2$ for W as follows:

- $\boldsymbol{q}_1 = \boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- $\boldsymbol{q}_2 = \boldsymbol{w}_2 - \left(\begin{array}{c} \text{projection of} \\ \boldsymbol{w}_2 \text{ onto } \boldsymbol{q}_1 \end{array} \right) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ -4 \\ 2 \end{bmatrix}$

  Note. $\boldsymbol{q}_2$ is the error of the projection of $\boldsymbol{w}_2$ onto $\boldsymbol{q}_1$. This guarantees that it is orthogonal to $\boldsymbol{q}_1$. On the other hand, since $\boldsymbol{q}_2$ is a combination of $\boldsymbol{w}_2$ and $\boldsymbol{q}_1$, we know that $\boldsymbol{q}_2$ actually is in W.

We have thus found the orthogonal basis $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{2}{3} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ for W (if we like, we can, of course, drop that $\frac{2}{3}$).

Important comment. By normalizing, we get an orthonormal basis for W: $\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

Practical comment. When implementing Gram–Schmidt on a computer, it is beneficial (slightly less work) to normalize each $\boldsymbol{q}_i$ during the Gram–Schmidt process. This typically introduces square roots, which is why normalizing at the end is usually preferable when working by hand.

Comment. There are, of course, many orthogonal bases $\boldsymbol{q}_1, \boldsymbol{q}_2$ for W. Up to the length of the vectors, ours is the unique one with the property that $\text{span}\{\boldsymbol{q}_1\} = \text{span}\{\boldsymbol{w}_1\}$ and $\text{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2\} = \text{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$.

A matrix Q has orthonormal columns $\iff$ $Q^T Q = I$

Why? Let $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ be the columns of Q. By the way matrix multiplication works, the entries of $Q^T Q$ are dot products of these columns:

$$\begin{bmatrix} - & \boldsymbol{q}_1^T & - \\ - & \boldsymbol{q}_2^T & - \\ & \vdots & \end{bmatrix} \begin{bmatrix} | & | & \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{bmatrix}$$

Hence, $Q^T Q = I$ if and only if $\boldsymbol{q}_i^T \boldsymbol{q}_j = 0$ (that is, the columns are orthogonal), for $i \neq j$, and $\boldsymbol{q}_i^T \boldsymbol{q}_i = 1$ (that is, the columns are normalized).

Example 71. $Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix}$ obtained from Example 70 satisfies $Q^T Q = I$.

Armin Straub
straub@southalabama.edu